RLP-ALG1A · RLP Algebra 1A — Data, Linear Equations and Systems

Unit 4: Two-Variable Statistics

Practice Workbook. The same problems, in the same order, as the Practice sections of the online lessons.

Name _____ Period _____

Lesson 4.1 — Two-Way Tables: Is There an Association?

FLUENCY

1. A survey records two yes/no questions. Yes to both: 10; yes to A only: 55; yes to B only: 16; no to both: 20. Find P(A), P(A and B), P(A | B), and decide whether A and B look independent.

2. A survey records two yes/no questions. Yes to both: 53; yes to A only: 18; yes to B only: 59; no to both: 21. Find P(A), P(A and B), P(A | B), and decide whether A and B look independent.

3. A survey records two yes/no questions. Yes to both: 29; yes to A only: 24; yes to B only: 14; no to both: 36. Find P(A), P(A and B), P(A | B), and decide whether A and B look independent.

4. A survey records two yes/no questions. Yes to both: 43; yes to A only: 14; yes to B only: 15; no to both: 18. Find P(A), P(A and B), P(A | B), and decide whether A and B look independent.

5. A table of 120 students: grade 9 — bus 18, car 30, walk 12; grade 12 — bus 10, car 42, walk 8. Find P(car | grade 12), P(grade 12 | car) and the joint relative frequency of grade 12 and car.

6. In the same table, what percent of all students walk?

APPLICATION

7. In a survey of 200 Missouri students (illustrative), 72 of 110 urban students and 38 of 90 rural students use a public library weekly. Compare conditional relative frequencies and decide whether the variables are associated.

8. In the same survey, what percent of all students use a library weekly?

9. A table shows 50 ninth graders and 70 twelfth graders. If there were no association, and 30% of each grade had a job, how many in each grade would have jobs?

CHALLENGE

10. Of 400 students, 25% are in group B, and 40% of group B answered yes to question A. How many answered yes and are in B? If A and B were independent, how many students overall would answer yes?

REVIEW

11. Solve the system: $6x + 5y = 31$ and $3x - 5y = -22$

12. Find the slope of the line through $(-3, -6)$ and $(0, 2)$.

Lesson 4.2 — Scatter Plots and Lines of Fit

FLUENCY

1. Use technology to find the least-squares line and the correlation coefficient for (2, 11), (6, 21), (7, 20), (13, 36), (14, 36).

2. Use technology to find the least-squares line and the correlation coefficient for (1, 6), (3, 5), (10, 18), (11, 19), (14, 22).

3. Use technology to find the least-squares line and the correlation coefficient for (7, -12), (10, -14), (11, -17), (13, -16), (14, -19).

4. Use technology to find the least-squares line and the correlation coefficient for (4, 11), (5, 12), (6, 18), (8, 25), (10, 27), (12, 37), (13, 35).

5. Use technology to find the least-squares line and the correlation coefficient for (1, -2), (2, -2), (9, 3), (12, 7), (13, 7), (14, 5).

6. Using predicted arm span ≈ 1.021 × height − 3.757, predict the arm span for heights 150, 170 and 190 cm. Which prediction is extrapolation?

APPLICATION

7. Six students report hours studied and test scores: (1, 62), (2, 68), (3, 71), (4, 75), (5, 82), (6, 85). Find the least-squares line and predict the score for 4.5 hours.

8. In the study model, interpret the slope and the intercept in context.

9. A model for Missouri counties (illustrative) is predicted median rent = 12.5 × (median income in thousands) + 600. Interpret 12.5.

CHALLENGE

10. The arm-span data have mean height 167 cm and mean arm span 166.75 cm. Show that the least-squares line passes through (mean height, mean arm span).

REVIEW

11. Write the equation, in slope-intercept form, of the line through (−6, 9) and (0, 3).

12. Solve: $-4x + 6 = 9x + 32$

Lesson 4.3 — Residuals and Correlation: How Good Is the Fit?

FLUENCY

1. Use technology to find the least-squares line and the correlation coefficient for (3, 19), (4, 18), (5, 27), (7, 32), (8, 36), (9, 38), (11, 42).

2. Use technology to find the least-squares line and the correlation coefficient for (1, 9), (2, 6), (3, 5), (6, -1), (7, -1), (10, -2), (11, -3).

3. Use technology to find the least-squares line and the correlation coefficient for (4, -4), (8, 2), (9, 1), (11, 3), (12, 5).

4. Use technology to find the least-squares line and the correlation coefficient for (1, 0), (3, -2), (4, 0), (5, 0), (11, 8), (12, 6).

5. For the model $\hat{y} = 2x + 5$, find the residuals for (3, 12), (5, 14) and (8, 22).

6. A residual is −3. Is the point above or below the line, and did the model overestimate or underestimate?

APPLICATION

7. Find the residual for the student 168 cm tall with a 170 cm arm span in Lesson 4.2.

8. Order these by strength of linear association: $r = 0.45$, $r = -0.92$, $r = 0.78$, $r = -0.10$.

9. A residual plot shows residuals positive at both ends and negative in the middle. What does that say about the linear model?

CHALLENGE

10. For the points (1, 2), (2, 4), (3, 5), find the least-squares line and r, and show that the residuals add to zero.

REVIEW

11. Solve the system: $-5x + 4y = 0$ and $6x - 2y = -14$

12. Solve: $|x + 3| = 5$

Lesson 4.4 — Correlation Is Not Causation: Screens, Sleep and Hidden Variables

FLUENCY

1. Use technology to find the least-squares line and the correlation coefficient for (1, 7), (3, 17), (5, 23), (6, 22), (10, 33), (11, 41), (12, 41).

2. Use technology to find the least-squares line and the correlation coefficient for (7, 21), (9, 29), (10, 30), (12, 36), (13, 42), (14, 45).

3. Use technology to find the least-squares line and the correlation coefficient for (4, 18), (5, 20), (7, 22), (9, 29), (10, 27), (11, 31), (14, 38).

4. Rewrite "Each hour of screens causes 0.64 hour less sleep" in association language.

5. Is a study that surveys students about breakfast and grades observational or experimental?

6. What does random assignment do in an experiment?

APPLICATION

7. Use predicted sleep ≈ 8.85 − 0.64x to predict sleep for 2.5 hours of late screen time.

8. Across lake resort towns, daily ice cream sales and sunburn cases are strongly correlated. Name a confounding variable.

9. Design an experiment to test whether a 9 p.m. phone curfew increases sleep.

CHALLENGE

10. Use the sleep model to predict sleep for 14 hours of late screen time, and explain why the prediction is not trustworthy.

REVIEW

11. Solve and graph on a number line: $4x - 6 \geq -14$

12. For the data 7, 8, 11, 17, 18, 37, 39: find the mean, median, range and interquartile range. (Find the quartiles as the medians of the lower and upper halves, leaving out the overall median when the count is odd.)

Unit 4 checkpoint

1. Of 60 ninth graders, 12 have jobs. What is the conditional relative frequency of having a job given grade 9?
   A. 20%
   B. 10%
   C. 26.7%
   D. 37.5%
2. Of 45 students with jobs, 33 are twelfth graders: _____% (to one decimal).
3. True or false: Equal conditional relative frequencies across rows suggest no association.
4. The slope of predicted arm span ≈ 1.02 × height − 3.76 means…
   A. about 1.02 cm more arm span per extra cm of height, on average
   B. every student's arm span is 1.02 times height
   C. arm span is −3.76 at birth
   D. r = 1.02
5. A residual is actual minus _____.
6. Which r shows the strongest linear association?
   A. −0.95
   B. 0.8
   C. 0.5
   D. 0
7. True or false: r = 0.9 means 90% of the points lie on the line.
8. A residual plot curves. What should you conclude?
   A. A linear model is the wrong form
   B. The model is excellent
   C. r must be 0
   D. There are no residuals
9. With predicted sleep ≈ 8.85 − 0.64x, the prediction for x = 3 is about _____ hours.
10. True or false: A strong correlation in observational data proves causation.

11. What establishes cause most convincingly?
    A. A randomized experiment
    B. A strong r
    C. A large sample survey
    D. A steep slope

12. Predicting arm span for a 100 cm child from data on 152-183 cm students is…
    A. extrapolation
    B. interpolation
    C. a residual
    D. a correlation